Molecular epidemiology of vaccine preventable viral diseases in poultry

NKFIH-OTKA, grant no. 120201

Final report

Molecular epidemiology of infectious diseases combines the collection of genetic data of infectious agent(s) and the epidemiological background information linked to the case(s). In this project we aimed at studying the molecular epidemiology of vaccine preventable viral poultry diseases. We placed whole genome sequencing of viral agents in the focus of investigation. In the research proposal we offered three work packages (WP1-WP3); of these, the most important aims of WP1 and WP2 have been successfully attained.

In WP1 we made efforts to optimize sample processing in order to maximize the quality and quantity of NGS sequence reads. This was a complex task given that viruses under investigation belonged to different virus families characterized by various genomic structure (i.e. dsDNA, ssDNA, dsRNA or ssRNA; circular, linear, segmented) and various size (approx. 2 kb to 300 kb). Our ultimate goal to develop a universal virus nucleic acid (NA) enrichment protocol remained only desire. The following protocol was evaluated for several viruses: we used a protocol to start with filtration of viral particles through 0.45 uM filters followed by digestion of RNA and DNA molecules not protected by protein coat (i.e. ideally the majority of host origin nucleic acid) and then extracting the viral NA on filter columns. This approach resulted in up to 90% elimination of host origin NA molecules. The amount of viral genomic RNA/DNA obtained by this method was sufficient to assemble complete genomes, or, at least the coding regions at high sequencing coverage. We successfully applied this method to enrich the genomic NA of avian coronaviruses, paramyxoviruses, reoviruses, and adenoviruses (although enrichment of adenovirus DNA was not always needed).

This approach, however, was useless for a variety of high-morbidity poultry viruses, including poxviruses, polyomaviruses, some small ssRNA viruses (such as duck hepatitis A virus) and, occasionally, in case of influenza A virus and infectious bursal disease virus. The amount of viral genomic nucleic acids in clinical specimens or tissue cultures was likely one factor that has led to the failure of the above-mentioned viral NA enrichment protocol. To overcome this issue, for influenza A viruses we adapted a published, single primer pair based whole-genome amplifying protocol. The amplified cDNA was suitable for assembly of H5, H7 and H9 influenza virus genomes. In addition, a whole genome amplifying PCR was developed for infectious bursal disease virus. For duck hepatitis A virus we followed target-specific amplification of genomic RNA. In this protocol we amplified partially overlapping genomic RNA fragments (size ~2-3 kb). The amplified cDNA was suitable for NGS and genome assembly of these viruses. For poxviruses, we attempted several methods (including enrichment by capture oligo probes) but only a modified free-NA-digesting protocol has led to success. As collaborative partners we originally described this method for another large dsDNA virus, African swine fever virus. The published method was successfully applied when analyzing several avipoxvirus isolates.

In WP2 the aims were at the characterization of various field virus strains detected in vaccinated/non-vaccinated poultry flocks. In this WP we relied partly on the strain collections from past epizootics. The NÉBIH ÁDI (National Food Chain Safety Office, Veterinary Diagnostic Directorate) and CEVA Phylaxia had huge amount of virus isolates not only from Hungary but also from other countries in the world. Part of this WP was the identification of potential revertant or altered vaccine strains. Most live vaccines contain attenuated infectious viruses and various mechanisms (incl. point mutations, recombination) could, theoretically, lead to the emergence of revertant vaccine strains or a modified, vaccine-derived strain containing genomic fragments from a field strain by reassortment or recombination.

Orthomyxoviruses. Our partner at NEBIH ÁDI selected 30 subtype H5 avian influenza virus isolates from the recent bird flu epidemics (year 2020) for genome sequencing and involved our group in data generation and analysis. The majority of isolates collected in the southern region of Hungary showed marked genetic conservation in the H5 HA gene, the N8 NA gene and the backbone genes whereas another two strains belonged to slightly different genetic clusters. Data are currently compared with past H5 avian influenza outbreaks in Hungary as well as the ongoing epizootics in 2021.

Paramyxoviruses. We have selected 40 avian paramyxovirus (AAV1) strains collected from 2002-2016 and following a single passage on embryonated chicken eggs we amplified the viral genomes by random RT-PCR. We identified multiple genotypes among AAV1 strains having circulated in Hungary over the past decade. Genotype I and II strains were identified in domestic poultry, whereas genotype VI and XXI strains were detected in pigeon and dove species. Moreover, the 30 Hungarian genotype VI AAV1 strains could be classified into multiple sub-genotypes.

We determined the genomic sequence of a Newcastle disease virus (NDV) line obtained directly from the first NDV isolate, named Herts'33. This strain shared ≤ 90% nucleotide sequence identity with the NDV sequences available in the GenBank database, and formed a distinct branch in a phylogenetic tree. This branch may be considered to represent a separate NDV genotype. Our study indicates that investigation of the genomic sequences of old NDV strains that originated from the early outbreaks of Newcastle disease may alter the phylogenetic grouping of the NDV strains and provide data on the evolution of viral genomes over time.

We sequenced the genome of three avian paramyxovirus 2 strains and developed an RT-qPCR assay to detect them. Although avian paramyxovirus 2 is not a major pathogen of domestic poultry and vaccines are not available to control its spread, data indicate that the virus may be common in parts of the world, including northern Africa. The assay we developed may elucidate the true geographic distribution of these viruses and their possible etiologic role in poultry disease(s).

Reoviruses. Avian reoviruses are well-known pathogens seriously affecting the productivity of poultry industry. Reoviruses of chicken are known as major causes of tenosynovitis. Other disease forms of reovirus infections have also been described and reovirus is frequently isolated from asymptomatic chickens. A collection of reovirus strains was subjected to whole genome sequencing. For classification purposes we processed so far only the sigmaC coding open reading frame; these analyses showed that the strains in our collection belong to four major genetic clusters. The genetic distances within each cluster were considerable (up to 30-40%, pending on cluster). The low sequence similarity was seen not only in the nucleotide but also in the amino acid alignments, a finding that suggests significant antigenic divergence within each cluster. In collaboration with colleagues at CEVA Phylaxia who provided the majority of strains for genetic analysis, the basis of a new classification system is being elaborated.

Additional genetic diversity of reoviruses was seen in waterfowls and game birds. First, the complete genomic sequences of two orthoreovirus strains, D2533/4/1-10 and D2533/6/1-10, isolated from Pekin ducklings in Germany were determined. Pairwise sequence comparisons and phylogenetic analyses indicated that strain D2533/4/1-10 might have acquired its genomic segments from different strains, such as classical and novel waterfowl reoviruses, and a yet unknown orthoreovirus strain. D2533/6/1-10 proved to be only distantly related to previously described orthoreoviruses. Subsequently, a reovirus strains closely related to D2533/6/1-10 was identified in a Hungarian goose flock affected by decreased egg production. None of these viruses are closely related to an established orthoreovirus species, therefore they may represent a novel species that could have originated from a wild bird.

When analyzing pheasant origin reovirus strains we demonstrated that reoviruses of pheasants are of diverse origin. The complete or coding-complete genomic sequences of two Hungarian reovirus strains, D1996/2/1 and Reo/HUN/Pheasant/216/2015, were determined in a study. The strain D1996/2/1 was isolated in 2012 from birds with gizzard erosion, whereas the other strain was isolated in 2015 from

diarrheic pheasant poults. Phylogenetic analyses showed that none of the Hungarian isolates shared common origin with a pheasant reovirus detected recently in the United States. Additionally, we found that Reo/HUN/Pheasant/216/2015 is a multi-reassortant reovirus within the species avian orthoreovirus that shared genetic relationship with turkey, partridge and chicken reoviruses in the respective gene phylogenies, whereas two genes (μB and μNS) did not reveal any possible common ancestors. The other isolate, D1996/2/1, was found to be distantly related to previously described reoviruses raising the possibility that it might represent a novel orthoreovirus species or a new genogroup within the newly accepted species, Neoavian orthoreovirus.

Adenoviruses. The genomic analysis of 30 fowl adenovirus A (FAdV-A) revealed sequence conservation in the genome. The genome-wide distance was less than one percent. Yet, sophisticated bioinformatics tools detected evidence of recombination and at least 10 recombination patterns were seen among strains isolated from chicken in Europe, Asia and the Americas.

We also determined the genomic sequence of a Ukrainian strain of fowl adenovirus B (FAdV-B). The isolate (D2453/1) shared 97.2% to 98.4% nucleotide sequence identity with other viruses belonging to the species fowl adenovirus B. Marked genetic divergence was seen in the hexon, fiber, and ORF19 genes, and phylogenetic analysis suggested that recombination events had occurred in these regions. Our analysis revealed mosaicism in the recombination patterns, a finding that has also been described in the genomes of strains of FAdV-D and FAdV-E. The shared recombination breakpoints, affecting the same genomic regions in viruses belonging to different species, suggest that similar selection mechanisms are acting on the key neutralization antigens and epitopes in viruses of different FAdV species.

The preliminary sequence analysis of 90 randomly selected and fully sequenced FAdV-E isolates from a global collection classified them into serotype 8a or 8b. Among 45 FAdV-D, serotype 2 and 11 strains were the most common serotypes. In addition, we identified a few FAdV-D strains with recombination in the hexon gene; we plan to investigate whether this recombination event may influence the serotype specificity of the affected strains.

Coronaviruses. Infectious bronchitis (IB) of chicken is a high morbidity and mortality viral disease affecting the poultry industry worldwide. We sequenced IBV isolates from various geographic locations and years. Initially, 110 IBV isolates were selected for whole gneome sequencing by using the random RT-PCR protocol. With this approach we obtained that 12% of samples did not contain sufficient amount of IBV genomic sequence reads for genome assembly. Another 8% had no IBV reads at all questioning the initial diagnosis of IBV-related disease or suggesting the overgrow of IBV in cell culture by a competitive cytopathogenic virus; indeed these samples contained large amounts of sequence reads derived from other pathogenic poultry viruses, such as avian influenza virus, astrovirus, adenovirus, parvovirus, avulavirus, reovirus and retrovirus. We used a subset of the fully sequenced IBV genomes to study the importance of recombination in genomic diversification and potential interactions between field and vaccine strains. First we assembled the genome of 20 randomly selected strains from seven European countries. After sequencing, we created a genome sequence data set that contained 36 European origin field isolates and 33 vaccine strains. When analyzing these 69 IBV genome sequences, we identified 215 recombination events highlighting that some strains had multiple recombination breaking points. Recombination hot spots were identified mostly in the regions coding for non-structural proteins, and multiple recombination hot spots were identified in the nsp2, nsp3, nsp8, and nsp12 coding regions. Recombination occurred among different IBV genotypes and involved both field and vaccine IBV strains. Ninety percent of field strains and nearly half of vaccine strains showed evidence of recombination.

Analysis of additional IBV strains identified putative new genotypes in Africa and the Middle East. The putative new IBV genotype in Middle East evolved through recombination affecting the structure of the S protein coding gene.

Polyomaviruses. We initiated a sequencing study of goose hemorrhagic polyomavirus and related polyomaviruses – in these analyses we determined the evolutionary rates and inferred phylodynamics and phylogeography of goose haemorrhagic polyomavirus (GHPV; Anser anser polyomavirus 1). GHPV infects the internal organs causing haemorrhagic nephritis and enteritis of geese that may be fatal for goslings. In a partly retrospective study GHPV positive samples were selected from goose and duck samples. Samples were collected in Hungary between 2005 and 2019. In this period, 384 of the investigated 1,111 specimens were diagnosed as GHPV-positive by PCR assay. Twenty-two GHPV genomes were sequenced and subjected to phylogenetic and evolutionary analysis. Based on the sequence data, the mean evolutionary rates were estimated 6.57 × 10-6 -5.82 × 10-5 s/s/y for both GHPV complete genomes and individual genes, with negative selection acting on each gene.

To better understand the evolution of avian polyomaviruses, we investigated genomic evolution and selection constraint acting on other virus species. Our analyses suggested that GHPV evolves more slowly than budgerigar fledgling disease virus (BFDV), finch polyomavirus (FPyV) and canary polyomavirus (CaPyV) (1.39 × 10-4 s/s/y, 2.63 × 10-4 s/s/y and 1.41 × 10-4 s/s/y mean evolutionary rate, respectively). In general, purifying selection seems to act on the protein coding regions of APyV genomes, although positive Darwinian selection was also predicted in a few positions (e.g., in the large tumor antigen coding region of BFDV and GHPV and in the capsid protein sequences of BFDV). Currently, we do not understand the importance of these aa changes.

CRESS DNA viruses. Circoviruses and other replication-associated protein-encoding single stranded (CRESS) DNA viruses have been detected in a variety of animal taxa. In a study, cloacal swab samples (n = 90) were examined for CRESS DNA viruses from 31 wild bird species living at various aquatic sites in Hungary to identify possible reservoirs of viruses pathogenic to domestic poultry. A total of 30 (33.3%) specimens tested positive with pan-CRESS DNA virus specific PCR. Goose circovirus (GoCV), Duck associated cyclovirus 1 (DuACyV-1) and Garrulus glandarius associated circular virus 1 (GgaCV-1) were detected in nine, three and two different bird species, respectively. Selected specimens were subjected to whole genome sequencing. The obtained sequence data revealed conserved gene structure within the identified virus species and detected homologous (within GoCV) and possible heterologous recombination (within DuACyV-1) events. A novel genotype of goose circovirus was also identified. Our results provide new information on the genomic diversity and evolution of selected CRESS DNA viruses.

Picornaviruses. We analyzed the genomic diversity of duck hepatitis A virus. Using available GenBank records in the analysis and genome sequences we determined from the epizootics that occurred in mid-2000s in Hungary, we found evidence of intergenotype recombination events. Recombination affected DHAV-1 and DHAV-3 sequences. For example, a 750 nt long fragment of the 5' genomic region of three strains collected in China showed high identity with DHAV-1 sequences (minor parent), while the remaining 7 kb long genomic region aligned with DHAV-3 sequences (major parent). The recombination affected the VP0 capsid encoding region of the polyprotein in these strains. Selection constraint was calculated for the DHAV-1 coding sequences. The overall dN/dS ratio was low for the polyprotein region and single genes (≤0.159) but positive selection was predicted to act at some aa sites, primarily in the capsid coding VP3 and VP1 regions. The rates for single genes were relatively high with values in the range of 5.6286x10-4 – 1.1147x10-3, as typical for ssRNA viruses. The polymerase encoding 3D region showed the lowest values.

Birnaviruses. Reassortant strains of infectious bursal disease virus (IBDV) were detected in commercial broiler flocks in western Europe and in layers and organic broilers in northern Europe during 2017-2019. As collaborating laboratory, we participated in the genome sequencing of selected strains. Genetic analysis, based on hypervariable region of VP2 gene showed grouping together with very virulent IBDV strains (vvIBDV, Genogroup 3). The VP1 gene of these isolates was most closely related to D78, an attenuated IBDV strain. A recently described reassortant IBDV strain from Poland with similar genomic constellation (segment A from vvIBDV, segment B from attenuated strain) retained its pathogenicity

(80 % mortality in SPF chickens). Infection with the North-West European reassortant IBDVs described in this study showed subclinical feature in the field (without complicating agents) and when tested under standardized pathogenicity test in SPF layer chickens (no mortality or clinical signs, but marked bursa atrophy was observed). Diagnostic investigations suggested that co-infection with fowl adenovirus or chicken infectious anaemia virus exaggerated the outcome of the IBDV infection (10-20 % mortality).

Orthobunyaviruses. Our research group helped identify possible viral etiology in South-East Asian disease outbreaks characterized by high morbidity and mortality renal failure among broilers. Although similar clinical picture is commonly seen with astrovirus and infectious bronchitis virus infections, in the outbreak cases we identified a novel orthobunyavirus (related to the mosquito-borne Umbre virus, Turlock serogroup). We also developed a quantitative RT-PCR assay to detect these viruses. With this assay we provided evidence that the virus was active in consecutive years mainly during the rainy season when mosquito activity is the greatest. In experimental infections the isolated orthobunyavirus reproduced the disease seen in the field fulfilling the Koch postulates. We also found evidence that the virus persists in the spleen for several weeks in survivors following recovery. Unfortunately, the paper we published from the new results was subsequently retracted. The official explanation was as follows: "…We assert that research was performed in good faith, that all the experimental data contained in the article are well founded and scientifically valid, and that there was no scientific misconduct. However, we subsequently were made aware of further information about the epidemiologic and clinical observations made locally in Malaysia, which brings into question the geographic location where the noted virus originated..."

Poxviruses. Three poxvirus strains (one chicken, one turkey and one pigeon origin strain) were available for analysis. Avian poxviruses have large genomic DNA and it is challenging to determine their genomic sequence even if they are isolated. We encountered the same issue and tried to overcome this by optimizing the enrichment of viral genomic DNA. The obtained consensus sequences covered the near full length genome and shared >99% identity with the genome of reference chicken, turkey and pigeon origin poxvirus strains in GenBank.

Conclusions

Whole genome sequencing of viruses directly from clinical samples without prior isolation or amplification step is currently challenging. Evolving target NA enrichment (e.g., capture oligo based enrichment) protocols may serve as good alternatives by increasing the clinical sensitivity of NGS-based techniques in virus diagnostics and molecular epidemiological studies. We made NGS technology to our collaborating partners working in the field of poultry health readily available. We improved NGS-based protocols for a variety of poultry viruses. As viruses we studied represented at least 12 virus families we anticipate that the 'know-how's we created over the past years form a solid basis to method adaptation for other animal viruses. More importantly, by analyzing hundreds of viral genomic sequences we made several new discoveries. These included the detection of numerous new viral variants and description of molecular evolutionary mechanisms of selected viruses neglected in the past.

Published papers in the topics (* indicates where the principal investigator was first, last and/or corresponding author)

Farkas SL, Varga-Kugler R, Marton S, Lengyel G, Palya V, Bányai K*. Genomic sequence and phylogenetic analyses of two novel orthoreovirus strains isolated from Pekin ducks in 2014 in Germany. Virus Res. 2018;257:57-62 Q2 IF: 2,736

[Palya V, Kovács EW, Marton S, Tatár-Kis T, Felföldi B, Forró B, Domán M, Bányai K*. Novel Orthobunyavirus Causing Severe Kidney Disease in Broiler Chickens, Malaysia, 2014-2017. Emerg Infect Dis. 2019;25:1110-1117 (Retraction in Emerg Infect Dis. 2020;26:1964.)]

Fehér E, Bálint Á, Marton S, Bali K, Belák S, Bányai K*. Coding-complete genome sequencing suggests that Newcastle disease virus challenge strain Herts'33 (IVMP) may represent a distinct genotype. Arch Virol. 2020;165:245-248. Q1 IF: 2,243

Kaszab E, Lengyel G, Marton S, Dán Á, Bányai K*, Fehér E. Occurrence and genetic diversity of CRESS DNA viruses in wild birds: a Hungarian study. Scientific Reports 2020;10:7036 Q1/D1 IF: 3,998

Kaszab E, Marton S, Dán Á, Farsang A, Bálint Á, Bányai K*, Fehér E. Molecular epidemiology and phylodynamics of goose hemorrhagic polyomavirus. Transboundary and Emerging Diseases 2020;67:2602-2608 Q1/D1 IF: 4,188

Mató T, Tatár-Kis T, Felföldi B, Jansson DS, Homonnay Z, Bányai K, Palya V. Occurrence and spread of a reassortant very virulent genotype of Infectious Bursal Disease Virus with altered VP2 amino acid profile and pathogenicity in some European countries. Veterinary Microbiology 2020;245:108663 Q1/D1 IF: 3,03

Kaszab E, Marton S, Erdélyi K, Bányai K, Fehér E. Genomic evolution of avian polyomaviruses with a focus on budgerigar fledgling disease virus. Infect Genet Evol. 2021;90,104762 Q1 IF: 2,773

Homonnay Z, Jakab S, Bali K, Kaszab E, Kiss I, Palya V, Bányai K*. Genome sequencing of a novel variant of Fowl adenovirus B reveals mosaicism in the pattern of homologous recombination events. Arch Virol. 2021;in press Q1 IF: 2,243

Farkas S, Varga-Kugler R, Ihász K, Marton S, Gál J, Palya V, Bányai K*. Genomic characterization of avian and neoavian orthoreoviruses detected in pheasants. Virus Res. 2021;in press Q2 IF: 2,934

Bali K, Bálint Á, Farsang A, Marton S, Nagy B, Kaszab E, Belák S, Palya V, Bányai K*. Recombination events shape the genomic evolution of infectious bronchitis virus in Europe. Viruses. 2021;in press Q1 IF: 3,816

Kiss I, Homonnay ZG, Mató T, Bányai K, Palya V. Research Note: An overview on distribution of fowl adenoviruses. Poultry Sci. 2021; in press Q1/D1 IF: 2,659